CLAIMS

1.   A method comprising:

identifying an initial set of pitch value candidates within each frame of a plurality of frames of received audio content utilizing a first pitch estimation algorithm; and

reducing the initial set of pitch value candidates to a select set of pitch value candidates based, at least in part, on pitch value re-scoring utilizing a second pitch estimation algorithm, wherein the select set of pitch values are selected in substantially real-time.

2.   The method according to claim 1, further comprising:

calculating a transition probability between at least one of the select pitch value candidates of adjacent frames.

3.   The method according to claim 2, further comprising:

selecting a pitch value within each frame with the highest transition probability between adjacent frames as the pitch value for the frame.

4.   The method according to claim 2, wherein the transition probability is based, at least in part, on dynamic programming configured to determine a significantly best path between different pitch candidates of adjacent frames.

5.   The method according to claim 2, further comprising:

smoothing a curve representing the select pitch values over a plurality of frames based, at least in part, on other information.

6. The method according to claim 5, wherein other information includes one or more of an energy value for each frame, a zero crossing rate of the audio content, and/or a vocal tract spectrum of the audio content.

7. The method according to claim 1, wherein identifying the initial set of pitch value candidates within each frame comprises:

passing each frame of audio content through an average magnitude difference function (AMDF); and

selecting N near-zero minima pitch values in the audio content as the initial set of pitch value candidates.

8. The method according to claim 7, wherein N is set to 288 pitch value candidates, selected as the initial set of pitch value candidates based, at least in part, on the AMDF.

9. The method according to claim 1, wherein identifying a select set of pitch values comprises:

generating a local score for each of the initial set of pitch value candidates utilizing a normalized cross-correlation function (NCCF); and

selecting M pitch value candidates with the highest local score.

10. The computer readable media having computer instructions for performing acts comprising:

identifying an initial set of pitch value candidates within each frame of a plurality of frames of received audio content utilizing a first pitch estimation algorithm; and

reducing the initial set of pitch value candidates to a select set of pitch value candidates based, at least in part, on pitch value re-scoring utilizing a second pitch estimation algorithm, wherein the select set of pitch values are selected in substantially real-time.

11. The computer readable media according to claim 10, having further computer instructions for performing acts comprising:

calculating a transition probability between at least one of the select pitch value candidates of adjacent frames.

12. The computer readable media according to claim 11, having further computer instructions for performing acts comprising:

selecting a pitch value within each frame with the highest transition probability between adjacent frames as the pitch value for the frame.

13. The computer readable media according to claim 11, wherein the transition probability is based, at least in part, on dynamic programming configured to determine a significantly best path between different pitch candidates of adjacent frames.

14. The computer readable media according to claim 11, having further computer instructions for performing acts comprising:

smoothing a curve representing the select pitch values over a plurality of frames based, at least in part, on other information.

15. The computer readable media according to claim 14, wherein other information includes one or more of an energy value for each frame, a zero crossing rate of the audio content, and/or a vocal tract spectrum of the audio content.

16. The computer readable media according to claim 10, wherein identifying the initial set of pitch value candidates within each frame comprises:

passing each frame of audio content through an average magnitude difference function (AMDF); and

selecting N near-zero minima pitch values in the audio content as the initial set of pitch value candidates.

17. The computer readable media according to claim 16, wherein N is set to 288 pitch value candidates, selected as the initial set of pitch value candidates based, at least in part, on the AMDF.

18. The computer readable media according to claim 10, wherein identifying a select set of pitch values comprises:

generating a local score for each of the initial set of pitch value candidates utilizing a normalized cross-correlation function (NCCF); and

selecting M pitch value candidates with the highest local score.

19.   An apparatus comprising logic configured to receive audio content, identify an initial set of pitch value candidates within each frame of a plurality of frames of the received audio content utilizing a first pitch estimation algorithm, and reduce the initial set of pitch value candidates to a select set of pitch value candidates based, at least in part, on pitch value re-scoring utilizing a second pitch estimation algorithm, wherein the select set of pitch values are selected in substantially real-time.

20.   The apparatus according to claim 19, wherein the logic is further configured to calculate a transition probability between at least one of the select pitch value candidates of adjacent frames-.

21.   The apparatus according to claim 20, wherein the logic is further configured to select a pitch value within each frame with the highest transition probability between adjacent frames as the pitch value for the frame.

22.   The apparatus according to claim 20, wherein the transition probability is based, at least in part, on dynamic programming configured to determine a significantly best path between different pitch candidates of adjacent frames.

23.   The apparatus according to claim 20, wherein the logic is further configured to smoothing a curve representing the select pitch values over a plurality of frames based, at least in part, on other information.

24. The apparatus according to claim 23, wherein the other information includes one or more of an energy value for each frame, a zero crossing rate of the audio content, and/or a vocal tract spectrum of the audio content.

25. The apparatus according to claim 19, wherein, when the logic identifies the initial set of pitch value candidates within each frame, the logic is further configured to pass each frame of audio content through an average magnitude difference function (AMDF), and select N near-zero minima pitch values in the audio content as the initial set of pitch value candidates.

26. The apparatus according to claim 25, wherein N is set to 288 pitch value candidates, selected as the initial set of pitch value candidates based, at least in part, on the AMDF.

27. The apparatus according to claim 19, wherein, when the logic identifies the select set of pitch values, the logic is further configured to generate a local score for each of the initial set of pitch value candidates utilizing a normalized cross-correlation function (NCCF), and select M pitch value candidates with the highest local score.